



HFW16

RAW SEQUENCE LISTING

DATE: 08/17/2004

PATENT APPLICATION: US/10/049,515

TIME: 10:31:35

Input Set : A:\34536726.app

Output Set : N:\CRF4\08172004\J049515.raw

3 <110> APPLICANT: PLOWMAN, GREGORY D.
 4 WHYTE, DAVID
 5 MARTINEZ, RICARDO
 6 HILL, RONALD
 7 FLANAGAN, PETER
 8 LIOUBIN, MARIO
 10 <120> TITLE OF INVENTION: NOVEL PROTEIN PHOSPHATASES AND DIAGNOSIS AND TREATMENT
 11 OF PHOSPHATASE-RELATED DISORDERS
 13 <130> FILE REFERENCE: 034536-0726
 15 <140> CURRENT APPLICATION NUMBER: 10/049,515
 16 <141> CURRENT FILING DATE: 2002-06-14
 18 <150> PRIOR APPLICATION NUMBER: PCT/US00/22158
 19 <151> PRIOR FILING DATE: 2000-08-11
 21 <160> NUMBER OF SEQ ID NOS: 45
 23 <170> SOFTWARE: PatentIn Ver. 3.2
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 715
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Mus sp.
 30 <400> SEQUENCE: 1
 31 actccccca cgatgggagt gcaaccccc aacttctctt ggggtgcttc gggacggctg 60
 32 gccggactgg cgttgccccg gctgccgcg cactaccagt tcttgcctga ccagggtgtg 120
 33 cggcacctgg tgtccctgac ggagcgcgga cccctcaca gtgacagctg tcccggcctc 180
 34 acgctgcacc gaatgcgcat cctgacttt tgccgcgcgt ccccggaaca gatcgaccaa 240
 35 tttgtgaaga tcgtggacga ggccaatgcc cggggagagg ctgttgaggt gcaactgtgc 300
 36 ctaggctttg gccgcactgg caccatgcta gcttgcctat tggatgaagg gcggtgcttg 360
 37 gccgcaggag atgccattgc tgagatccgg cgcctgcgac caggatccat tgagacgtat 420
 38 gaacaggaga aggccgtctt ccagttctac cagcgaacaa aatgaggact tcaacaagcc 480
 39 cgcctttccc cctccccaac tcttgcggcc agggaggaag gggagtgaac taaagtactg 540
 40 cctccttcag gtccctctga ctctattgg acaaaagtag tcttcccca aagccataac 600
 41 gtggccggca ggatggcga gacccacaa aaatgaggta ataactgata agaactcatc 660
 42 accgtgcat agcatgtaca cagcactccc aatacatctg ggtggttgaa aagac 715
 45 <210> SEQ ID NO: 2
 46 <211> LENGTH: 150
 47 <212> TYPE: PRT
 48 <213> ORGANISM: Mus sp.
 50 <400> SEQUENCE: 2
 51 Met Gly Val Gln Pro Pro Asn Phe Ser Trp Val Leu Pro Gly Arg Leu
 52 1 5 10 15
 54 Ala Gly Leu Ala Leu Pro Arg Leu Pro Ala His Tyr Gln Phe Leu Leu
 55 20 25 30
 57 Asp Gln Gly Val Arg His Leu Val Ser Leu Thr Glu Arg Gly Pro Pro
 58 35 40 45

(pg.6)

ENTERED

RAW SEQUENCE LISTING

DATE: 08/17/2004

PATENT APPLICATION: US/10/049,515

TIME: 10:31:35

Input Set : A:\34536726.app

Output Set: N:\CRF4\08172004\J049515.raw

```

60 His Ser Asp Ser Cys Pro Gly Leu Thr Leu His Arg Met Arg Ile Pro
61      50                      55                      60
63 Asp Phe Cys Pro Pro Ser Pro Glu Gln Ile Asp Gln Phe Val Lys Ile
64 65                      70                      75                      80
66 Val Asp Glu Ala Asn Ala Arg Gly Glu Ala Val Gly Val His Cys Ala
67                      85                      90                      95
69 Leu Gly Phe Gly Arg Thr Gly Thr Met Leu Ala Cys Tyr Leu Val Lys
70                      100                      105                      110
72 Glu Arg Ala Leu Ala Ala Gly Asp Ala Ile Ala Glu Ile Arg Arg Leu
73                      115                      120                      125
75 Arg Pro Gly Ser Ile Glu Thr Tyr Glu Gln Glu Lys Ala Val Phe Gln
76 130                      135                      140
78 Phe Tyr Gln Arg Thr Lys
79 145                      150

```

82 <210> SEQ ID NO: 3

83 <211> LENGTH: 968

84 <212> TYPE: DNA

85 <213> ORGANISM: Mus sp.

87 <400> SEQUENCE: 3

```

88 cgcaactcgc ggccccgtag atggccctag ggtcgcaggg gttatgggca cctcggagggc 60
89 tgcaccgccc cctttcgcgc gcgtcgcccc cgcgtcttc atcggaatg cgcgagccgc 120
90 ggggtgcgacg gagctgctgg tgcgcgcggg catcactttg tgcgtcaatg tctcccgcca 180
91 gcagcccggg ccgcgcgcgc ccggagtggc ggaactacgc gtaccctgtg tgcagacccc 240
92 agctgaggac ctgctgacac acctggagcc cacctgtgccc gccatggaag ccgcggtgcg 300
93 cgacggcggc tctgtctcg tgtactgcaa gaacggccgc agtcgctcag ccgccgtctg 360
94 caccgcctac ctaatgcggc accgcggcca cagcctggat cgcgccttcc agatggtgaa 420
95 gacgcgccgc ccggtagccg agcccaattt ggggttctgg gctcagctgc agaagtacga 480
96 gcagaccctt caggcccagg ccactcctgcc ccgggagccc attgatccgg agtaagccga 540
97 ctgttcggct gctgggtgac caagcgtcta tactgaaagg aagtgtccct tccctccttt 600
98 ttctattagg cagctggctt tgggtgttgc cccatcttga tggtagtaca ggaacgtcta 660
99 ctgagtagga ggacttcgtt tattcatcat gtttgacca aatccaaacc agcacgtttt 720
100 aggtagagaa attgagtga gtagtgtctg ggaagcctac gaacggttga tagcgagtga 780
101 tagatcagag tcttagctgc ctactccaag ggagtgcctg ggtttgtagg cagaacctat 840
102 ctgtctcctg aacttctggc cccttagaaa tgaacataga gtctcccagc aggagctcat 900
103 gggccacact ttggcttaca gctgctgtgc catggaaggg aggcctgca gactgcagct 960
104 gagcgact

```

107 <210> SEQ ID NO: 4

108 <211> LENGTH: 163

109 <212> TYPE: PRT

110 <213> ORGANISM: Mus sp.

112 <400> SEQUENCE: 4

```

113 Met Gly Thr Ser Glu Ala Ala Pro Pro Pro Phe Ala Arg Val Ala Pro
114 1      5      10      15
116 Ala Leu Phe Ile Gly Asn Ala Arg Ala Ala Gly Ala Thr Glu Leu Leu
117      20      25      30
119 Val Arg Ala Gly Ile Thr Leu Cys Val Asn Val Ser Arg Gln Gln Pro
120      35      40      45
122 Gly Pro Arg Ala Pro Gly Val Ala Glu Leu Arg Val Pro Val Phe Asp
123      50      55      60

```

RAW SEQUENCE LISTING

DATE: 08/17/2004

PATENT APPLICATION: US/10/049,515

TIME: 10:31:35

Input Set : A:\34536726.app

Output Set: N:\CRF4\08172004\J049515.raw

```

125 Asp Pro Ala Glu Asp Leu Leu Thr His Leu Glu Pro Thr Cys Ala Ala
126 65          70          75          80
128 Met Glu Ala Ala Val Arg Asp Gly Gly Ser Cys Leu Val Tyr Cys Lys
129          85          90          95
131 Asn Gly Arg Ser Arg Ser Ala Ala Val Cys Thr Ala Tyr Leu Met Arg
132          100          105          110
134 His Arg Gly His Ser Leu Asp Arg Ala Phe Gln Met Val Lys Ser Ala
135          115          120          125
137 Arg Pro Val Ala Glu Pro Asn Leu Gly Phe Trp Ala Gln Leu Gln Lys
138          130          135          140
140 Tyr Glu Gln Thr Leu Gln Ala Gln Ala Ile Leu Pro Arg Glu Pro Ile
141 145          150          155          160
143 Asp Pro Glu
146 <210> SEQ ID NO: 5
147 <211> LENGTH: 1067
148 <212> TYPE: DNA
149 <213> ORGANISM: Mus sp.
151 <400> SEQUENCE: 5
152 cccgcgtccg gcgagcgcgg agcggcgagc cccggcgggc catggggagt gggatgagcc 60
153 agatcctgcc gggcctgtac attggcaact tcaaagacgc aagagatgca gaacagttga 120
154 gcaggaacaa ggtgacacac attctttctg tgcacgatac tgccaggccc atgttgagg 180
155 gagttaaata cctgtgtatt ccagcggcag acacaccatc tcaaacctg acaagacatt 240
156 tcaaagaaag cattaaattc attcatgagt gccgactcca gggtgagagc tgtcttgtag 300
157 attgcctggc tggggtctcc aggagtgtga cattggtgat cgcatacatc atgactgtca 360
158 ccgacttttg ctgggaagat gccttgacac ctgttcgtgc ggggaggtcc tgtgccaaac 420
159 ccaacctggg ctttcaaagg cagctgcagg agtttgagaa acatgaagtg caccagtatc 480
160 ggcaatggct gagagaagag tatggagaga accctttgcg ggatgcagaa gaagccaaaa 540
161 atattctggc tgccccggga attctgaagt actgggcctt tctcagaaga ctgtaatgta 600
162 cctgaagttt ctgaaatatt gcaaagttca ggctgggtgct gccaaaaaga aaagtgatgt 660
163 aaagtttatt tttaagaatc caatagtgat ttgtataact gttttttttt catttttaac 720
164 caaatgcatg tataatcatg ttggaatatg ttaagatcta tggatattct gtagcaagag 780
165 aaaatatctt tgctttaatc ccactgctgt gggtgttctt tggacctgac cgatgctcat 840
166 acaataatct caagagccct gtctgtttcg taatagtaac tacttctcat gaacactacc 900
167 caaggaggaa gcctgcacct ggaagtgtg cagtgtgagc tctgccctcc tgttaagttc 960
168 tccagctcta gacatgtctc ttggtgtgtg ttttatctac tgggtgttatt ctatatggta 1020
169 gaattaccaa aagctattca gatttcttaa taaagggcaa atcaacc 1067
172 <210> SEQ ID NO: 6
173 <211> LENGTH: 184
174 <212> TYPE: PRT
175 <213> ORGANISM: Mus sp.
177 <400> SEQUENCE: 6
178 Met Gly Ser Gly Met Ser Gln Ile Leu Pro Gly Leu Tyr Ile Gly Asn
179 1          5          10          15
181 Phe Lys Asp Ala Arg Asp Ala Glu Gln Leu Ser Arg Asn Lys Val Thr
182          20          25          30
184 His Ile Leu Ser Val His Asp Thr Ala Arg Pro Met Leu Glu Gly Val
185          35          40          45
187 Lys Tyr Leu Cys Ile Pro Ala Ala Asp Thr Pro Ser Gln Asn Leu Thr
188          50          55          60

```

RAW SEQUENCE LISTING

DATE: 08/17/2004

PATENT APPLICATION: US/10/049,515

TIME: 10:31:35

Input Set : A:\34536726.app

Output Set: N:\CRF4\08172004\J049515.raw

```

190 Arg His Phe Lys Glu Ser Ile Lys Phe Ile His Glu Cys Arg Leu Gln
191   65              70              75              80
193 Gly Glu Ser Cys Leu Val His Cys Leu Ala Gly Val Ser Arg Ser Val
194              85              90              95
196 Thr Leu Val Ile Ala Tyr Ile Met Thr Val Thr Asp Phe Gly Trp Glu
197              100             105             110
199 Asp Ala Leu His Thr Val Arg Ala Gly Arg Ser Cys Ala Asn Pro Asn
200              115             120             125
202 Leu Gly Phe Gln Arg Gln Leu Gln Glu Phe Glu Lys His Glu Val His
203              130             135             140
205 Gln Tyr Arg Gln Trp Leu Arg Glu Glu Tyr Gly Glu Asn Pro Leu Arg
206 145              150             155             160
208 Asp Ala Glu Glu Ala Lys Asn Ile Leu Ala Ala Pro Gly Ile Leu Lys
209              165             170             175
211 Tyr Trp Ala Phe Leu Arg Arg Leu
212              180

```

215 <210> SEQ ID NO: 7

216 <211> LENGTH: 597

217 <212> TYPE: DNA

218 <213> ORGANISM: Mus sp.

220 <400> SEQUENCE: 7

```

221 atggactcgc tacagaagca ggaacttcgg aggccaaaga ttcattggggc agtccagggtg 60
222 tccccctacc agccaccac actggcctct ctgcagcgat tgctgtgggt ccgtcggact 120
223 gccacactga cccacatcaa tgagggtctg cccaaccttt tcttgggaga tgcgtatgct 180
224 gccagagaca agggtcgtct aatccagctg ggcattaccc atgttgtgaa tgtggctgcg 240
225 ggcaagttcc aggtggacac aggtgccaag ttctaccgtg gaacacctct ggagtactat 300
226 ggcatagagg ctgatgacaa ccccttcttt gacctcagcg tccactttct gcctgttgct 360
227 cgttacatca gagatgccct caatattccc cgaagccgag tgctgggtcca ctgcgctatg 420
228 ggggtgagtc gctctgccac aattgtcttg gccttctcca tgatcttcga gaacatgaca 480
229 ctggtagatg ccatccagac ggtgcaggcc caccgagata tctgtcccaa ctcaggcttc 540
230 ctccgacagc tccaggttct ggacaacagg ctgaggcggg aaacaggaag actctga 597

```

233 <210> SEQ ID NO: 8

234 <211> LENGTH: 198

235 <212> TYPE: PRT

236 <213> ORGANISM: Mus sp.

238 <400> SEQUENCE: 8

```

239 Met Asp Ser Leu Gln Lys Gln Glu Leu Arg Arg Pro Lys Ile His Gly
240   1              5              10              15
242 Ala Val Gln Val Ser Pro Tyr Gln Pro Pro Thr Leu Ala Ser Leu Gln
243              20              25              30
245 Arg Leu Leu Trp Val Arg Arg Thr Ala Thr Leu Thr His Ile Asn Glu
246              35              40              45
248 Val Trp Pro Asn Leu Phe Leu Gly Asp Ala Tyr Ala Ala Arg Asp Lys
249              50              55              60
251 Gly Arg Leu Ile Gln Leu Gly Ile Thr His Val Val Asn Val Ala Ala
252   65              70              75              80
254 Gly Lys Phe Gln Val Asp Thr Gly Ala Lys Phe Tyr Arg Gly Thr Pro
255              85              90              95
257 Leu Glu Tyr Tyr Gly Ile Glu Ala Asp Asp Asn Pro Phe Phe Asp Leu

```

RAW SEQUENCE LISTING

DATE: 08/17/2004

PATENT APPLICATION: US/10/049,515

TIME: 10:31:35

Input Set : A:\34536726.app

Output Set: N:\CRF4\08172004\J049515.raw

```

258          100          105          110
260 Ser Val His Phe Leu Pro Val Ala Arg Tyr Ile Arg Asp Ala Leu Asn
261          115          120          125
263 Ile Pro Arg Ser Arg Val Leu Val His Cys Ala Met Gly Val Ser Arg
264          130          135          140
266 Ser Ala Thr Ile Val Leu Ala Phe Leu Met Ile Phe Glu Asn Met Thr
267 145          150          155          160
269 Leu Val Asp Ala Ile Gln Thr Val Gln Ala His Arg Asp Ile Cys Pro
270          165          170          175
272 Asn Ser Gly Phe Leu Arg Gln Leu Gln Val Leu Asp Asn Arg Leu Arg
273          180          185          190
275 Arg Glu Thr Gly Arg Leu
276          195
279 <210> SEQ ID NO: 9
280 <211> LENGTH: 828
281 <212> TYPE: DNA
282 <213> ORGANISM: Mus sp.
284 <400> SEQUENCE: 9
285 atgcactccc tgaaccaaga aatcaaagca ttctcccggg ataatctcag gaagcagtgc 60
286 accagggtga ccacgctaac tggaaagaaa cttatagaaa cctgggaaga tgccacagtt 120
287 catgttgttg agacagagcc cagcggtggg ggtggttgtg gctatgtgca ggaccttacc 180
288 ttggacctgc aagttggcgt tattaagccc tggttgcttc tggggtcaca ggatgctgct 240
289 catgacctgg agctactgag aaagcataag gtgactcata ttctcaatgt tgcatatgga 300
290 gttgaaaatg ctttcctcag tgagttttaca tataagacca tttctatact ggatgtgcct 360
291 gaaaccaata tcctgtctta ttttcagaa tgttttgagt ttattgagca agcaaaactg 420
292 aaggatggcg tggttctcgt gcactgtaat gcagggtgtt ccagggtgct tgcaattgtc 480
293 attggcttcc tcatgagttc tgaagaagcc actttcacca ctgccctgtc gttggtgaaa 540
294 gaggcgagac catccatatg tccgaatcct ggcttcattg aacaactccg caccctacca 600
295 gtaggcaagg agagcaatgg aggtgacaaa gtgcccgcgg aggacacgac cgggtggtctg 660
296 tgatctgtac tccagcagag gcaaacgact tctgcatcag actctgtcct cttgcctgtg 720
297 ttttggaagg aaacttgga aacttccttt ttctgtgttc ttttaccagt ggaaatgaac 780
298 gttacctttg tcgccctgaa ttaatacatt ttaaagtttt accttttc 828
301 <210> SEQ ID NO: 10
302 <211> LENGTH: 220
303 <212> TYPE: PRT
304 <213> ORGANISM: Mus sp.
306 <400> SEQUENCE: 10
307 Met His Ser Leu Asn Gln Glu Ile Lys Ala Phe Ser Arg Asp Asn Leu
308 1 5 10 15
310 Arg Lys Gln Cys Thr Arg Val Thr Thr Leu Thr Gly Lys Lys Leu Ile
311 20 25 30
313 Glu Thr Trp Glu Asp Ala Thr Val His Val Val Glu Thr Glu Pro Ser
314 35 40 45
316 Gly Gly Gly Gly Cys Gly Tyr Val Gln Asp Leu Thr Leu Asp Leu Gln
317 50 55 60
319 Val Gly Val Ile Lys Pro Trp Leu Leu Leu Gly Ser Gln Asp Ala Ala
320 65 70 75 80
322 His Asp Leu Glu Leu Leu Arg Lys His Lys Val Thr His Ile Leu Asn
323 85 90 95

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/049,515

DATE: 08/17/2004
TIME: 10:31:36

Input Set : A:\34536726.app
Output Set: N:\CRF4\08172004\J049515.raw

Please Note:

se of n and/or Xaa have been detected in the Sequence Listing. Please review the
equence Listing to ensure that a corresponding explanation is presented in the <220>
o <223> fields of each sequence which presents at least one n or Xaa.

eq#:11; N Pos. 39
eq#:29; N Pos. 1954,2010,2032
eq#:41; N Pos. 57

VERIFICATION SUMMARY

PATENT APPLICATION: **US/10/049,515**

DATE: 08/17/2004

TIME: 10:31:36

Input Set : **A:\34536726.app**

Output Set: **N:\CRF4\08172004\J049515.raw**

L:361 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:1093 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:1920
L:1094 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:1980
L:1881 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0